

FIG. 1

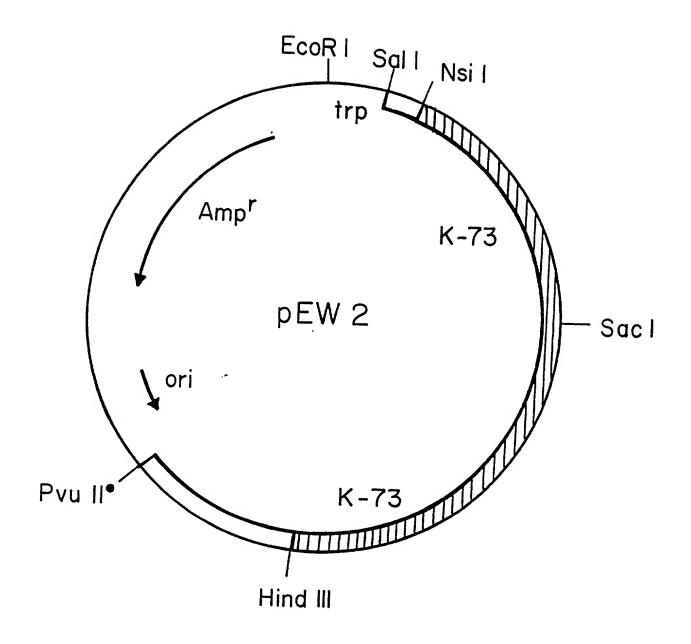


FIG. 2

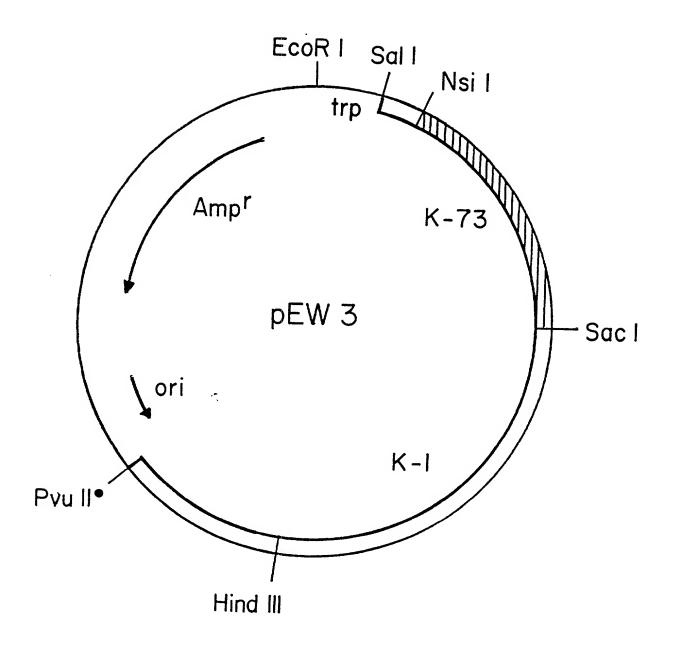


FIG. 3

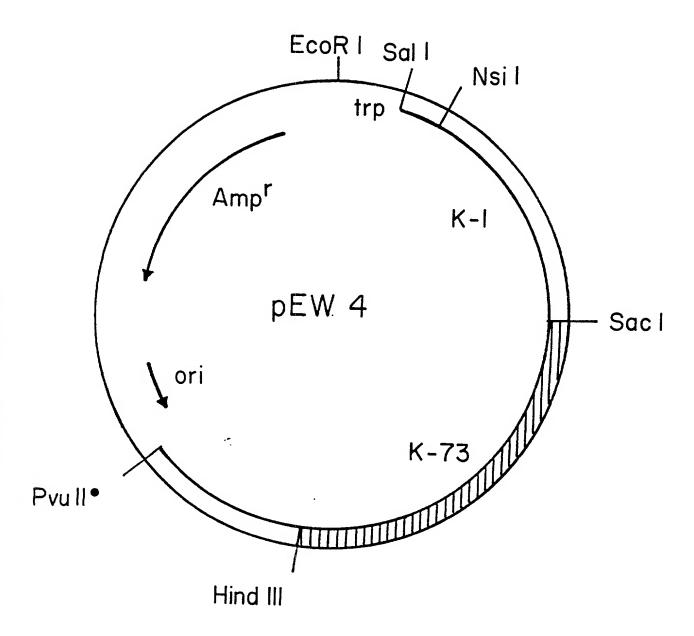


FIG. 4

	(start	HD-73)	ATG	GATAACAATC	400
		CCTTATAATT		CCCTGAAGTA	
GAAGTATTAG	GTGGAGAAAG	AATAGAAACT	GGTTACACCC	CAATCGATAT	500
TTCCTTGTCG			TGAATTTGTT	CCCGGTGCTG	•
GATTTGTGTT	AGGACTAGTT	GATATAATAT	GGGGAATTTT	TGGTCCCTCT	600
CAATGGGACG	CATTTCTTGT	ACAAATTGAA	CAGTTAATTA	ACCAAAGAAT	
AGAAGAATTC	GCTAGGAACC	AAGCCATTTC	TAGATTAGAA	GGACTAAGCA	700
ATCTTTATCA	AATTTACGCA	GAATCTTTTA	GAGAGTGGGA	AGCAGATCCT	
ACTAATCCAG	CATTAAGAGA	AGAGATGCGT	ATTCAATTCA	ATGACATGAA	800
CAGTGCCCTT	ACAACCGCTA	TTCCTCTTTT	TGCAGTTCAA	AATTATCAAG	
TTCCTCTTTT	ATCAGTATAT	GTTCAAGCTG	CAAATTTACA	TTTATCAGTT	900
TTGAGAGATG	TTTCAGTGTT	TGGACAAAGG	TGGGGATTTG	ATGCCGCGAC	
TATCAATAGT	CGTTATAATG	ATTTAACTAG	GCTTATTGGC	AACTATACAG	1000
ATTATECTET	ACGCTGGTAC	AATACGGGAT	TAGAACGTGT	ATGGGGACCG	
GATTCTAGAG	ATTEGETAAG	GTATAATCAA	TTTAGAAGAG	AATTAACACT	1100
AACTGTATTA	GATATCGTTG	CTCTGTTCCC	GAATTATGAT	AGTAGAAGAT	
ATCCAATTCG	AACAGTTTCC	CAATTAACAA	GAGAAATTTA	TACAAACCCA	1200
GTATTAGAAA	ATTTTGATGG	TAGTTTTCGA	GGCTCGGCTC	AGGGCATAGA	
AAGAAGTATT	AGGAGTCCAC	ATTTGATGGA	TATACTTAAC	AGTATAACCA	1300
TCTATACGGA	TECTCATAGE	GGTTATTATT	ATTGGTCAGG	GCATCAAATA	
ATGGCTTCTC	CTGTAGGGTT	TTCGGGGCCA		TTCCGCTATA	1400
TGGAACTATG	GGAAATGCAG	CTCCACAACA	ACGTATTGTT	GCTCAACTAG	
GTCAGGGCGT	GTATAGAACA	TTATCGTCCA	CTTTATATAG	AAGACCTTTT	1500
AATATAGGGA	TAAATAATCA	ACAACTATCT		GGACAGAATT	
TGCTTATGGA	ACCTCCTCAA	ATTTGCCATC	CGCTGTATAC		1600
GAACGGTAGA	TTCGCTGGAT	GAAATACCGC		CAACGTGCCA	
CCTAGGCAAG	GATTTAGTCA	TCGATTAAGC	CATGTTTCAA	TGTTTCGTTC	1700
AGGCTTTAGT	AATAGTAGTG	TAAGTATAAT			
(start	HD-1)	CCAACGT	TTTCTTGGCA	GCATCGCAGT	1900
GCTGAATTTA	ATAATATAAT	TEETTEATEA	CAAATTACAC	AAATACCTTT	
AACAAAATCT	ACTAATCTTG	GCTCTGGAAC	TTCTGTCGTT	AAAGGACCAG	2000
GATTTACAGG	AGGAGATATT	CTTCGAAGAA	CTTCACCTGG	CCAGATTTCA	
	TAAATATTAC		TCACAAAGAT		2100
AATTCGCTAC	GCTTCTACTA	CAAATTTACA			
GAAGACCTAT	TAATCAGGGT	AATTTTTCAG	CAACTATGAG	TAGTGGGAGT	2200
AATTTACAGT		TAGGACTGTA	GGTTTTACTA	CTCCGTTTAA	
CTTTTCAAAT		TATTTACGTT		GTCTTCAATT	2300
CAGGCAATGA	AGTTTATATA	GATCGAATTG	AATTTGTTCC	GGCAGAAGTA	
ACCTTTGAGG	CAGAATATGA	TTTAGAAAGA	GCACAAAAGG	CGGTGAATGA	2400
GCTGTTTACT	TCTTCCAATC	AAATCGGGTT	AAAAACAGAT	GTGACGGATT	
ATCATATTGA	TCAAGTATCC	AATTTAGTTG	AGTGTTTATC	AGATGAATTT	2500
TGTCTGGATG	AAAAACAAGA	ATTETCCGAG	AAAGTCAAAC	ATECGAAGE	2500
ACTTAGTGAT	GAGCGGAATT	TACTTCAAGA	TCCAGACTTC	AGAGGGATCA	2600
ATAGACAACT	AGACCGTGGC	TGGAGAGGAA	GTACGGATAT	TARRATREA	2000
GGAGGCGATG	ACGTATTCAA	AGAGAATTAC	GTTACGCTAT	TRESTACETT	2700
TGATGAGTGC	TATCCAACGT	ATTTATATCA	AAAAATAGAT	GARTERADAT	-/

FIG. 5A

TAAAAGCCTA TACCCGTTAT CAATTAAGAG GGTATATCGA AGATAGTCAA 2800 GACTTAGAAA TCTATTTAAT TCGCTACAAT GCAAAACATG AAACAGTAAA TGTGCCAGGT ACGGGTTCCT TATGGCCGCT TTCAGCCCAA AGTCCAATCG 2900 GAAAGTGTGG AGAGCCGAAT CGATGCGCGC CACACCTTGA ATGGAATCCT GACTTAGATT GTTCGTGTAG GGATGGAGAA AAGTGTGCCC ATCATTCGCA 3000 TCATTTCTCC TTAGACATTG ATGTAGGATG TACAGACTTA AATGAGGACC TAGGTGTATG GGTGATCTTT AAGATTAAGA CGCAAGATGG GCACGCAAGA 3100 CTAGGGAATC TAGAGTTTCT CGAAGAGAAA CCATTAGTAG GAGAAGCGCT AGCTCGTGTG AAAAGAGCGG AGAAAAAATG GAGAGACAAA CGTGAAAAAT 3200 TGGAATGGGA AACAAATATC GTTTATAAAG AGGCAAAAGA ATCTGTAGAT GCTTTATTTG TAAACTCTCA ATATGATCAA TTACAAGCGG ATACGAATAT 3300 TGCCATGATT CATGCGGCAG ATAAACGTGT TCATAGCATT CGAGAAGCTT ATCTGCCTGA GCTGTCTGTG ATTCCGGGTG TCAATGCGGC TATTTTTGAA 3400 GAATTAGAAG GGCGTATTTT CACTGCATTC TCCCTATATG ATGCGAGAAA TETCATTAAA AATGGTGATT TTAATAATGG CTTATCCTGC TGGAACGTGA 3500 AAGGGCATGT AGATGTAGAA GAACAAAACA ACCAACGTTC GGTCCTTGTT CTTCCGGAAT GGGAAGCAGA AGTGTCACAA GAAGTTCGTG TCTGTCCGGG 3600 TCGTGGCTAT ATCCTTCGTG TCACAGCGTA CAAGGAGGGA TATGGAGAAG GTTGCGTAAC CATTCATGAG ATCGAGAACA ATACAGACGA ACTGAAGTTT 3700 AGCAACTGCG TAGAAGAGGA AATCTATCCA AATAACACGG TAACGTGTAA TGATTATACT GTAAATCAAG AAGAATACGG AGGTGCGTAC ACTTCTCGTA 3800 ATCGAGGATA TAACGAAGCT CCTTCCGTAC CAGCTGATTA TGCGTCAGTC TATGAAGAAA AATCGTATAC AGATGGACGA AGAGAGAATC CTTGTGAATT 3900 TAACAGAGGG TATAGGGATT ACACGCCACT ACCAGTTGGT TATGTGACAA AAGAATTAGA ATACTTCCCA GAAACCGATA AGGTATGGAT TGAGATTGGA 4000 GAAACGGAAG GAACATTTAT CGTGGACAGC GTGGAATTAC TCCTTATGGA GGAA (end HD-1)

FIG. 5B

MDNNPNINECIPYNCLSNPEVEVLGGERIE GΥ T P I D I SLSLTQFLLSEF VPGAGFVLGL V D I IWGIFGPSQWDAFLVQIEQLINQR FARNQAI SRLEGLSNLYQIYAE SFREWE TNPALREEMRIQFNDMNSALTTAIPLF V P L L S V Y V Q A A N L H L S V L R D V S WGF DAA \mathbf{T} INSRYNDLTRLIGN ΥT YNTGLE RVWG ΡD SRDWVRYNQFRREL T L LDIVALFPNYDSRRYP IRTVSQLT RΕ Ι VLE NFDG SFRGSAQG IERSIRSP HLMD Ι TDAHRGYYYWSGHQIMAS PVG F T FPLYGTMGNAAPQORIVAOLGOGV TLYRRPFNIGINNQQLSVLDGT E SNLPSAVYRKSGTVDSLDEI Ρ P PRQGFSHRLSHVSMFRSGFSNS S V S ΙI F QHRSAE FNNIIPSSQITQ ΙP L T VVKGPGF T S T GGDILRRT SPGO Ι T TAPLSQRYRVRIRYASTTNL QF Η INQGNFSATM SSGSNLQS GSFR Т Т FNF SNGSSVF TLSAHVFNSGNE V Y RIEFVPAEV TFEAEYDLERAOKAVNEL SNQIGLKTDVTDYHIDQVSNLVECLSD LDEKQELSEKVKHAKRLSDERNLLQD GINRQLDRGWRGSTDITIQGGDDVFKE TLLGTFDECÝPTYLYQKIDESKLKA LRGYIEDSQDLEIYLIRYNAKHETVN V GT LWPL SAQSPIGKCGEPNRCAPHLE WNPD C S C R EKCAHHSHHFSLDIDVG DG C TD L N EDLGVWVIFKIKTQDGHARLGNLE F L E ELVGEALARVKRAEKKWRDKREKLE W Ε Т SVDALFVNSQYDQLQAD YKEAKE TNI AADKRVHS IREAYLPELSV ΙP GVNAA LEGRI F TAFSLYDARNVIKNGDFNN GLS CW NVKGHVDVE EQNNQRSVLVLP Ε W Α E V SOE GRGYILRVTAYKEGYG VRVCP Ε G C V T I Η Ε Ι ENNTD ELKF SNCVEEE Ι ΥP NN T V T C N D Y V YGGAYTSRNRGYNEAPSVPADYAS Y V KSYTDGRRENPCEFNRGYRDYT P L PVGY TKELEYFPETDKVWIEIGETEGTFIVDSV ELLLMEE

FIG. 6

(c+:	ert HD-1)	ATE	• ***		-
		TTTAAGTAAC	ATAACAATCO	, GAALAICAAI	
TEGAGAAAA	ATAGAAACTE	: IIIAAGIAAC	AATCGATATT	TOCTTOTOGO	
TAACGCAATT					
	ATATAATATG				
	CAAATTGAAC				
CTAGGAACCA					
	AATCTTTTAG				
ATTAAGAGAA				CTAATCCAGC	
CAACCGCTAT	· · - · · · -			TCCTCTTTTA	1000
TCAGTATATG		***************************************	TTATCAGTTT	TGAGAGATGT	
TTCAGTGTTT					1100
GTTATAATGA					
	ATACGGGATT				1200
TTGGGTAAGG	TATAATCAAT	TTAGAAGAGA			
ATATCGTTGC	TCTATTCTCA		GTCGAAGGTA		1300
ACAGTTTCCC	AATTAACAAG			TATTAGAAAA	
TTTTGATGGT		GAATGGCTCA		CAGAATATTA	1400
GGCAACCACA		ATCCTTAATA	GTATAACCAT	TTATACTGAT	
	GCTTTAATTA	TTGGTCAGGG		CAGCTTCTCC	1500
TGTAGGGTTT			CCCTTTATTT	GGGAATGCGG	
GGAATGCAGC		CTTGTCTCAT	TAACTGGTTT	GGGGATTTTT	1600
AGAACATTAT	,	ATATAGAAGA	ATTATACTTG	GTTCAGGCCC	
AAATAATCAG				TCTTTTGCCT	1700
CCCTAACGAC	CAACTTGCCT	TCCACTATAT	ATAGACAAAG		
GATTCACTAG	ATGTAATACC	GCCACAGGAT	AATAGTGTAC		1800
		GTCATGTTAC			
GAGCAGTITA	CACCTTGAGA		(stop HD-1)		
000700700	(start			ATGTTCTCTT	
	TAGTGCTGAA				1800
	CTGCAGTGAA		CTTTTTAATG	GTTCTGTAAT	
	GGATTTACTG	_ · · ·	AGTTAGATTA	AATAGTAGTG	1900
GAAATAACAT			AAGTTCCAAT	TCACTTCCCA	
TCGACATCTA		AGTTCGTGTA			2000
		GGGGTAATTC		TCCAATACAG	
				TGATTTTGGT	2100
TATTTTGAAA		TTTTACATCT			
		CTGCAGGAGT	GATAATAGAC	AGATTTGAAT	2200
TTATTCCAGT	TACTGCAACA	CTCGAGGCTG	AATATAATCT	BOBABAABB	

	CAGAAGGCGG			ACAAACCAAC	TAGGGCTAAA	2300
	AACAAATGTA			AGTGTCCAAT	TTAGTTACGT	
	ATTTATCGGA			AGCGAGAATT	GTCCGAGAAA	2400
	GTCAAACATG			CGCAATTTAC	TCCAAGATTC	
	AAATTTCAAA			ACGTGGGTGG	GGCGGAAGTA	2500
	CAGGGATTAC			TATTTAAAGA	AAATTACGTC	
	ACACTATCAG	GTACCTTTGA		CCAACATATT	TGTATCAAAA	2600
	AATCGATGAA	TCAAAATTAA	AAGCCTTTAC	CCGTTATCAA	TTAAGAGGGT	
	ATATCGAAGA	TAGTCAAGAC	TTAGAAATCT	ATTTAATTCG	CTACAATGCA	2700
	AAACATGAAA	CAGTAAATGT	GCCAGGTACG	GGTTCCTTAT	GGCCGCTTTC	
	AGCCCAAAGT	CCAATCGGAA	AGTGTGGAGA	GCCGAATCGA		2800
	ACCTTGAATG	GAATCCTGAC	TTAGATTGTT	CGTGTAGGGA		
	TGTGCCCATC	ATTCGCATCA	TTTCTCCTTA	GACATTGATG	TAGGATGTAC	2900
	AGACTTAAAT	GAGGACCTAG	GTGTATGGGT	GATCTTTAAG	ATTAAGACGC	
	AAGATGGGCA	CGCAAGACTA	GGGAATCTAG	AGTTTCTCGA	AGAGAAACCA	3000
	TTAGTAGGAG	AAGCGCTAGC	TCGTGTGAAA	AGAGCGGAGA	AAAAATGGAG	
	AGACAAACGT	GAAAAATTGG	AATGGGAAAC	AAATATCGTT		3100
	CAAAAGAATC	TGTAGATGCT	TTATTTGTAA	ACTCTCAATA	TGATCAATTA	
-	CAAGCGGATA	CGAATATTGC	CATGATTCAT	GCGGCAGATA	AACGTGTTCA	3200
•	TAGCATTCGA	GAAGCTTATC	TGCCTGAGCT	GTCTGTGATT	CCGGGTGTCA	
4	ATGCGGCTAT	TTTTGAAGAA	TTAGAAGGGC	GTATTTTCAC	TGCATTCTCC	3300
1	CTATATGATG	DGAGAAATGT	CATTAAAAAT	GGTGATTTTA	ATAATGGCTT	
•	ATCCTGCTGG	AACGTGAAAG	GGCATGTAGA	TGTAGAAGAA	CAAAACAACC	3400
1	AACGTTCGGT	CCTTGTTGTT	CCGGAATGGG	AAGCAGAAGT	GTCACAAGAA	
(STTCGTGTCT	GTCCGGGTCG	TGGCTATATC	CTTCGTGTCA	CAGCGTACAA	3500
(3GAGGGATAT	GGAGAAGGTT	GCGTAACCAT	TCATGAGATC	GAGAACAATA	
(CAGACGAACT	GAAGTTTAGC	AACTGCGTAG	AAGAGGAAAT	CTATCCAAAT	3600
f	AACACGGTAA	CGTGTAATGA	TTATACTGTA	AATCAAGAAG	AATACGGAGG	
	rgcgtacact	TCTCGTAATC	GAGGATATAA	CGAAGCTCCT	TCCGTACCAG	3700
E	CTGATTATGC	GTCAGTCTAT	GAAGAAAAAT	CGTATACAGA	TGGACGAAGA	
E	BAGAATCCTT	GTGAATTTAA	CAGAGGGTAT	AGGGATTACA	CGCCACTACC	3800
f	AGTTGGTTAT	GTGACAAAAG	AATTAGAATA	CTTCCCAGAA	ACCGATAAGG	
7	FATGGATTGA	GATTGGAGAA	ACGGAAGGAA		GGACAGCGTG	3900
6	SAATTACTCC	TTATGGAGGA	A (end HD-7	' 3),	_	

FIG. 7B

MDNNPNINECIPYNCLSNPEVEVLGGERIE IDISLSLTQFLLSEFVPGAGFVLGL Ρ IWGIFGPSQWDAFPVQIEQLINQRIEE FARNQAISRLEGLSNLYQIYAESFRE TNPALREEMRIQFNDMNSALTTAIPLLAV QNYQVPLLSVYVQAANLHLSVLRD V S DAATINSRYNDLTRLIGNYT DΥ YNTGLE G P D S R D W V R Y N Q F R R RVWIVALFS NYDSRRYPIRTVSQLT R ENFDG SFRGMAQRIEQNIRQPHLMD DVHRGFNYWSGHQITAS ΙY \mathbf{T} Ρ V PLFGNAGNAAPPVLVSLTGLG F Ι PLYRRIILGSGPNNQELFVL D P S IYRQRGTVDSLDV TNL ${f T}$ IPPODNS RAGF SHRLSHVTMLSQAAGAVY TLRAO SWI HRSAEFNNIIASDSIT QΙ PAVK S VISGPGF T GGDLVRLNSSGN IHF P S Y \mathbf{E} V Ρ T S TRYRVRVR YASV N WG N S S IFSNT VPAT АТ SLDNLO GY F SANAFT SSLGNIVGVRNF SG TAGV F I PVTATLEAEYNLERAQKAVN QLGLKTNVTDYHIDQVSNLVTYL F C L D E K R E L S E K V K H A K R L S D E R N L L Q F K D I N R Q P E R G W G G S T G I T I Q G G D D V F YVTLSGTFDECYPTYLYQKIDESKLKAF YQLRGYIEDSQDLE IYLIRYNAKHE V G S L W P L S A Q S P I G K C G E P N R C A P H L S CRDGEKCAHHSHHFSLD Ι D V LNEDLGVWVIFKIKTQDGHARL GNLE KPLVGE ALARVKRAEKKWRDKREKLE YKEAKE S V D A L F V N S Q Y D Q L Q A D T IHAADKRVHS IREAYLPELSVI Ρ GVN EELEGR Ι T A F SLYDARNVIKN G D WNVKGHVD V E E Q N N Q R S V L V V P E W Ε GRGYILRVTAYKEG YGE C V IENNTDELKFSNCVEEEIYPNN T V T C QE YGGAYTSRNRGYNEAPS \mathbf{E} V P YEEKSYTDGRRENPCEFNRGYRDY ΤP V YVTKELEYFPETDKVWIEIGETEGTFIVD SVELLLMEE

FIG. 8

		HD-73)		GATAACAAT	
	A TGAATGCAT		T GTTTAAGTA		
	GTGGAGAAA		T GGTTACACC		T -500
	G CTAACGCAA		G TGAATTTGT	r cccggtgct	. E
GATTTGTGT		r gatataata	T GGGGAATTT	T TGGTCCCTC1	600
CAATGGGAC			A CAGTTAATTA	ACCAAAGAAT	
AGAAGAATT			C TAGATTAGAA	GGACTAAGCA	700
ATCTTTATC	A AATTTACGCA	GAATCTTTT/	A GAGAGTGGGA	AGCAGATCCT	
ACTAATCCA	S CATTAAGAGA	AGAGATGCG		ATGACATGA	
CAGTGCCCTT	ACAACCGCTA	TTCCTCTTT			
TTCCTCTTTT	ATCAGTATAT	GTTCAAGCT			
TTGAGAGAT	TTTCAGTGTT	TGGACAAAGG			
TATCAATAGT	CGTTATAATE	ATTTAACTA			
ATTATGCTGT	ACGCTGGTAC	AATACGGGAT	TAGAACGTGT	ATGGGGACCG	
GATTCTAGAG	ATTEGETAAG	GTATAATCAA			
AACTGTATTA	GATATCGTTG	CTCTGTTCCC	GAATTATGAT	AGTAGAAGAT	
ATCCAATTCG	AACAGTTTCC				1200
GTATTAGAAA	ATTTTGATGG	TAGTTTTCGA			
AAGAAGTATT	AGGAGTCCAC	ATTTGATGGA		AGTATAACCA	
TCTATACGGA	TGCTCATAGG	GGTTATTATT	ATTGGTCAGG		
ATGGCTTCTC	CTGTAGGGTT		GAATTCACTT		1400
TGGAACTATG	GGAAATGCAG	CTCCACAÁCA	ACGTATTGTT	GCTCAACTAG	
GTCAGGGCGT	GTATAGAACA		CTTTATATAG		1500
AATATAGGGA			GTTCTTGACG	GGACAGAATT	
TGCTTATGGA	ACCTCCTCAA	ATTTGCCATC	CGCTGTATAC	AGAAAAAGCG	16001
GAACGGTAGA		GAAATACCGC			
	GATTTAGTCA	TCGATTAAGC			1700
AGGCTTTAGT		TAAGTATAAT	AAGAGCT (e		2700
(start		CCAACGT			1900
	ATAATATAAT	TCCTTCATCA	CAAATTACAC	AAATACCTTT	
	ACTAATCTTG	GCTCTGGAAC	TTCTGTCGTT		2000
	AGGAGATATT		CTTCACCTGG	CCAGATTTCA	
ACCTTAAGAG	TAAATATTAC	TGCACCATTA	TCACAAAGAT	ATCGGGTAAG	2100
	GCTTCTACTA	CAAATTTACA	ATTCCATACA	TCAATTGACG	
GAAGACCTAT			CAACTATGAG	TAGTGGGAGT	2200
AATTTACAGT	CCGGAAGCTT	TAGGACTGTA	GGTTTTACTA	CTCCGTTTAA	
CTTTTCAAAT		TATTTACGTT	AAGTGCTCAT	GTCTTCAATT	2300
CAGGCAATGA	AGTTTATATA	GATCGAATTG	AATTTGTTCC	RECARAGETA	
ALL I I GAGG	CAGAATATGA	TTTAGAAAGA	GCACAAAAGG	CEETGAATEA	2400
GUIGITIACT	TCTTCCAATC	AAATCGGGTT	AAAAACAGAT	GTGACGGATT	
ATCATATTGA	TCAAGTATCC	AATTTAGTTG	AGTGTTTATC	AGATGAATTT	2500
161L166A16	AAAAACAAGA	ATTGTCCGAG	AAAGTCAAAC	ATGCGAAGCG	
ACTIAGTGAT	GAGCGGAATT	TACTTCAAGA	TCCAAACTTC	ARABBRATCA .	2600
ATAGACAACT	AGACCGTGGC	TGGAGAGGAA	GTACGGATAT	TACCATCCAA	
				···•	

GGAGGCGATG	ACGTATTCAA	AGAGAATTAC	GTTACGCTAT	TEGETACCTT	2700
TGATGAGTGC	TATCCAACGT	ATTTATATCA	AAAAATAGAT	GAGTCGAAAT	
TAAAAGCCTA	TACCCGTTAT	CAATTAAGAG	GGTATATCGA	AGATAGTCAA	2800
GACTTAGAAA	TCTATTTAAT	TEGETACAAT	GCAAAACATG	AAACAGTAAA	
TGTGCCAGGT	ACGGGTTCCT	TATGGCCGCT	TTCAGCCCAA	AGTCCAATCG	2900
GAAAGTGTGG	AGAGCCGAAT	CGATGCGCGC	CACACCTTGA	ATGGAATCCT	
GACTTAGATT	GTTCGTGTAG	GGATGGAGAA	AAGTGTGCCC	ATCATTCGCA	3000
TCATTTCTCC	TTAGACATTG	ATGTAGGATG	TACAGACTTA	AATGAGGACC	
TAGGTGTATG	GGTGATCTTT	AAGATTAAGA	CGCAAGATGG	GCACGCAAGA	3100
CTAGGGAATC	TABAGTTTCT	CGAAGAGAAA	CCATTAGTAG	GAGAAGCGCT	
AGCTCGTGTG	AAAAGAGCGG	AGAAAAAATG	GAGAGACAAA	CGTGAAAAAT	3200
TGGAATGGGA	AACAAATATC	GTTTATAAAG	AGGCAAAAGA	ATCTGTAGAT	
GCTTTATTTG	TAAACTCTCA	ATATGATCAA	TTACAAGCGG	ATACGAATAT	3300
TGCCATGATT	CATGCGGCAG	ATAAACGTGT	TCATAGCATT	CGAGAAGCTT	
ATCTGCCTGA	GCTGTCTGT.G	ATTCCGGGTG	TCAATGCGGC	TATTTTTGAA	3400
GAATTAGAAG	GGCGTATTTT	CACTGCATTC	TCCCTATATG	ATGCGAGAAA	
TETCATTAAA	AATGGT.GATT	TTAATAATGG	CTTATCCTGC	TGGAACGTGA	3500
AAGGGCATGT	AGATGTAGAA	GAACAAAACA	ACCAACGTTC	GGTCCTTGTT	
CTTCCGGAAT	GGGAAGCAGA	AGTGTCACAA	GAAGTTCGTG	TCTGTCCGGG	3600
TCGTGGCTAT	ATCCTTCGTG	TCACAGCGTA	CAAGGAGGGA	TATGGAGAAG	
GTTGCGTAAC	CATTCATGAG	ATCGAGAACA	ATACAGACGA	ACTGAAGTTT	3700
AGCAACTGCG	TAGAAGAGGA	AATCTATCCA	AATAACACGG	TAACGTGTAA	
TGATTATACT	GTAAATCAAG	AAGAATACGG	AGGTGCGTAC	ACTTCTCGTA	3800
ATCGAGGATA	TAACGAAGCT	CCTTCCGTAC	CAGCTGATTA	TGCGTCAGTC	
TATGAAGAAA	AATCGTATAC	AGATGGACGA	AGAGAGAATC	CTTGTGAATT	3900
TAACAGAGGG	TATAGGGATT	ACACGCCACT	ACCAGTTGGT	TATGTGACAA	
AAGAATTAGA	ATACTTCCCA	GAAACCGATA	AGGTATGGAT	TGAGATTGGA	4000
GAAACGGAAG	GAACATTTAT	CGTGGACAGC	GTGGAATTAC	TCCTTATGGA	
GGAA (end h	HD-1)				

FIG. 9B

MDNNPNINECIPYNCLSNPEVEVLGGERIE ΤP IDIS LSLTQFLLSEFVPGAG VDIIW GIFG PSQWDAFLVQIE QLINQRIE ARNQAI SRLEGLSNLYQIYAE S F R PALREE M R IQFNDMNSALTTA Ι P L A V V P L L S V Y V Q A A N L H L S V L D V S V DAA INSRYNDLT \mathbf{T} R L I G N Y T D Y YNTG LΕ R V W G P D SRD W V R YNQFRR E L L VALFP NYDSRRYP IRTVSQLT RE G SF R G SAQGIERS IRSP H L TDAHRGYYYWSGHQIMAS Y PLY MGNAAPQQRIVAQLGQGV GT Y R TLYRRPFNIGINNQQLSVLDGTE F SNLP AVYRKSGTVDSLNEIP S Ρ QNN PRQE F S H R L S H V S M F R S G F S N S S V S Ι Ι SW QHRSAEFNNIIPSSQITQIPL T VVKGPG T S \mathbf{T} F GGDILRRT SP G Q QRYRVRIRYAS RVNI TAPLS ТТ N L INQGNF Ρ SATM S SGS S NLO G S F R Т NF Ρ F SNGS SVFT LSAHVFNS G N V ΙE F V Ρ Α EVTFEAEYDLERAQKAVN ЕL SNQIGLKTDVTDYHI DQVSNLVE CLS D EKVKHAKRLSDERNLLQDP EKQELS GINRQLDRGWRGSTDITIQGGDDVFKE LGTFDECYPTYLYQKIDESKLKAY GYIE D S QDLE IYLIRYNAKH Ε T V N V SAQSPIGKCGEPNRCA РL PHLE W C R DGEKCAHHSHHFSLDI D V G C EDLGV W V IFKIKTQDGHARLGN L E LVGEALARVKRAEKKWRDKREKLE Ε W Т YKEAKESVD Α LFVNSQYDQLQAD T N ΙA AADKRVHSIREAYLPELSV ΙP G VNAA Ι F EGRI F Т A F SLYDARNVIKNG D F NN G S VE EQNNQRSVLVLP V D E W E Α Ε V S Р GRGY ILRVTÄYKE G ΥG E G V T I Ι ELKF S N C NN D VEE E IYPNN \mathbf{T} V T C N D V GAY TSRNRGYNEA Ρ S V P Α D ΥA EEKSYTDGRRENPCEFNRGYRDY Τ P L V G V T K E L E Y F P E T D K V W I E I G E T E G T F ELLLMEE

	(start	HD-73)	ATG	GATAACAATC	400
CGAACATCAA	TGAATGCATT	CCTTATAATT	GTTTAAGTAA	CCCTGAAGTA	
GAAGTATTAG	GTGGAGAAAG	AATAGAAACT	GGTTACACCC	CAATCGATAT	500
TTCCTTGTCG	CTAACGCAAT	TTCTTTTGAG	TGAATTTGTT	CCCGGTGCTG	
GATTTGTGTT	AGGACTAGTT	GATATAATAT	GGGGAATTTT	TEGTCCCTCT	600
CAATGGGACG	CATTTCTTGT	ACAAATTGAA	CAGTTAATTA	ACCAAAGAAT	
AGAAGAATTC	GCTAGGAACC	AAGCCATTTC	TAGATTAGAA	GGACTAAGCA	700
ATCTTTATCA	AATTTACGCA	BAATCTTTTA	GAGAGTGGGA	AGCAGATCCT	
ACTAATCCAG	CATTAAGAGA	AGAGATGCGT	ATTCAATTCA	ATGACATGAA	800
CAGTGCCCTT	ACAACCGCTA	TTCCTCTTTT	TGCAGTTCAA	AATTATCAAG	
TTCCTCTTTT	ATCAGTATAT	GTTCAAGCTG	CAAATTTACA	TTTATCAGTT	900
TTGAGAGATG	TTTCAGTGTT	TGGACAAAGG	TGGGGATTTG	ATGCCGCGAC	
TATCAATAGT	CGTTATAATG	ATTTAACTAG	GCTTATTGGC	AACTATACAG	1000
ATTATGCTGT	ACGCTGGTAC	AATACGGGAT	TAGAACGTGT	ATGGGGACCG	
GATTCTAGAG	ATTGGGTAAG	GTATAATCAA	TTTAGAAGAG	AATTAACACT	1100
	GATATCGTTG	CTCTGTTCCC			
ATCCAATTCG	AACAGTTTCC	CAATTAACAA		TACAAACCCA	1200
GTATTAGAAA	ATTTTGATGG	TAGTTTTCGA			
AAGAAGTATT	AGGAGTCCAC	ATTTGATGGA		AGTATAACCA	1300
	TECTCATAGE		ATTGGTCAGG	GCATCAAATA	
	CTGTAGGGTT				1400
TGGAACTATG	GGAAATGCAG	CTCCACAACA	ACGTATTGTT	GCTCAACTAG	
GTCAGGGCGT	GTATAGAACA	TTATEGTECA			1500
AATATAGGGA	TAAATAATCA	ACAACTATCT	GTTCTTGACG	GGACAGAATT	
TGCTTATGGA	ACCTCCTCAA	ATTTGCCATC	CGCTGTATAC		1600
GAACGGTAGA	TTCGCTGGAT	GAAATACCGC			
CCTAGGCAAG	GÁTTTAGTCA	TCGATTAAGC			1700
AGGCTTTAGT	AATAGTAGTG	TAAGTATAAT	AAGAGCT (er	nd hd-73)	
(start	HD-1)	CCAACGT	TTTCTTGGCA		1900
GCTGAATTTA	ATAATATAAT	TCCTTCATCA			
AACAAAATCT	ACTAATCTTG		TTCTGTCGTT		2000
GATTTACAGG	AGGAGATATT	CTTCGAAGAA			
ACCTTAAGAG	TAAATATTAC			ATCGGGTAAG	2100
AATTCGCTAC	GCTTCTACTA	CAAATTTACA			
GAAGACCTAT	TAATCAGGGT	AATTTTTCAG		TAGTGGGAGT	2200
AATTTACAGT	CCGGAAGCTT	TAGGACTGTA			
CTTTTCAAAT	GGATCAAGTG	TATTTACGTT		GTCTTCAATT	2300
CAGGCAATGA	AGTTTATATA				
		TTTAGAAAGA			2400
GCTGTTTACT	TCTTCCAATC	AAATCGGGTT	AAAAACAGAT	GTGACGGATT	
		AATTTAGTTG			2500
TETETERATE	4344464	ATTGTCCGAG	AVVETEVAVE	ATECCAACCC	2300
		TACTTCAAGA			2400
		TGGAGAGGAA			2000
		AGAGAATTAC			2700
TRATEARTEC	TATCCACCT	ATTTATATCA	AAAAATAGAT	CACTCCAAAT	2700
. unit unut uu	· · · · · · · · · · · · · · · · · · ·	ULLI WILLIAM	HUH!	GHG I CGAAA I	

FIG. 11A

TAAAACCCTA	TARROCTTAT	DAATTAACAD	COTATATODA		0000
TAAAAGCCTA		CAATTAAGAG		AGATAGTCAA	2800
GACTTAGAAA	TCTATTTAAT	TCGCTACAAT	GCAAAACATG	AAACAGTAAA	
TGTGCCAGGT	ACGGGTTCCT	TATGGCCGCT	TTCAGCCCAA	AGTCCAATCG	2900
GAAAGTGTGG	AGAGCCGAAT	CGATGCGCGC	CACACCTTGA	ATGGAATCCT	
GACTTAGATT	GTTCGTGTAG	GGATGGAGAA	AAGTGTGCCC	ATCATTCGCA	3000
TCATTTCTCC	TTAGACATTG	ATGTAGGATG	TACAGACTTA	AATGAGGACC	
TAGGTGTATG	GGTGATCTTT	AAGATTAAGA	CGCAAGATGG	GCACGCAAGA	3100
CTAGGGAATC	TAGAGTTTCT	CGAAGAGAAA	CCATTAGTAG	GAGAAGCGCT	
AGCTCGTGTG	AAAAGAGCGG	AGAAAAAATG	GAGAGACAAA	CGTGAAAAAT	3200
TGGAATGGGA	AACAAATATC	GTTTATAAAG	AGGCAAAAGA	ATCTGTAGAT	
GCTTTATTTG	TAAACTCTCA	ATATGATCAA	TTACAAGCGG	ATACGAATAT	3300
TGCCATGATT	CATGCGGCAG	ATAAACGTGT	TCATAGCATT	CGAGAAGCTT	
ATCTGCCTGA	GCTGTCTGT.G	ATTCCGGGTG	TCAATGCGGC	TATTTTTGAA	3400
GAATTAGAAG	GGCGTATTTT	CACTGCATTC	TCCCTATATG	ATGCGAGAAA	
TGTCATTAAA	AATGGTGATT	TTAATAATGG	CTTATCCTGC	TGGAACGTGA	3500
AAGGGCATGT	AGATGTAGAA	GAACAAAACA	ACCAACGTTC	GGTCCTTGTT	
CTTCCGGAAT	GGGAAGCAGA	AGTGTCACAA	GAAGTTCGTG	TCTGTCCGGG	3600
TEGTGGCTAT	ATCCTTCGTG	TCACAGCGTA	CAAGGAGGGA	TATGGAGAAG	
GTTGCGTAAC	CATTCATGAG	ATCGAGAACA	ATACAGACGA	ACTGAAGTTT	3700
AGCAACTGCG	TAGAAGAGGA	AATCTATCCA	AATAACACGG	TAACGTGTAA	
TGATTATACT	GTAAATCAAG	AAGAATACGG	AGGTGCGTAC	ACTTCTCGTA	3800
ATCGAGGATA	TAACGAAGCT	CCTTCCGTAC	CAGCTGATTA	TGCGTCAGTC	
TATGAAGAAA	AATCGTATAC	AGATGGACGA	AGAGAGAATC	CTTGTGAATT	3900
TAACAGAGGG	TATAGGGATT	ACACGCCACT	ACCAGTTGGT	TATGTGACAA	
AAGAATTAGA	ATACTTCCCA	GAAACCGATA	AGGTATGGAT	TGAGATTGGA	4000
GAAACGGAAG	GAACATTTAT	CGTGGACAGC		TCCTTATGGA	
GGAA (end)	HD-1)				

FIG. 11B

M D N N P N I N E C I P Y N C L S N P E V E V L G G E R I E TGYTPIDI SLSLTQFLLSEFVPGAGF V I F G P S Q W D A F L V Q I E Q L I N Q R I IIWG F A R N Q A I S R L E G L S N L Y Q I Y A E S F R E PTNPALREEMRIQFNDMNSALTTAI PLFAV QNYQVPLLSVYVQAANLHLSVLRD V RWGFDAATINSRYNDLTRLIGNYTD YAVRW YNTGLERVW G P D S R D W V R Y N Q F R R E L IVALFPNYDSRRYPIRTVSQLTREI PVLENFDGSFRGSAQGIEGSIRSPHLM IYTDAHKGEYYWSGHQIMAS Т ΡV LYGTMGNAAPQQRIVAQLGQGVYR T F Р TLYRRPFNIGINNQQLSVLD GTEF G T S S N L P S A V Y R K S G T V D S L D E ΙP PQNNNV PRQGFSHRLSHVSMFRSGFSN S S V S Ι IRA SWQHRSAEFNNIIPSSQITQIPLTK NLGSGTSVVKGPGF TGGDILRRT SPGO Ι TAPLSQRYRVRIRYASTTNLQFH LRVNI IDGRP INQGNFSATMSSGSNLQSGSFRT Т PFNFSNGSSVFTLSAHVFNS GNE RIEFVPAEVTFEAEYDLERAQKAVNEL F SNQIGLKTDVTDYHIDQVSNLVE CLSD LDEKQELSEKVKHAKRLSDERNLLQDP NF G I N R Q L D R G W R G S T D I T I Q G G D D V F K E TLLGTF DECYPTYLYQKIDESKLKAYT LRGYIEDSQDLEIYLIRYNAKHE TVN V G S L W P L S A Q S P I G K C G E P N R C A P H L E C S C R D G E K C A H H S H H F S L D I D V G Т EDLGVWVIFKIKTQDGHARLGNLE LVGEALARVKRAEKKWRDKREKLEW E TN YKEAKESVDALFVNSQYDQLQADTNIA AADKRVHSIREAYLPELSVIPGVNAA Ι EGRI SLYDARNVIKNGDFNN FTAF G L C VKGH VDVE EQNNQRSVLVLPEWEA GRGYILRVTAYKEGYGEG VRVCP C V T Ι Ι ENN DELKF SNCVEEEIYPNN T V Т C NQEE Y GGAYTSRNRGYNEAPS VPAD ΥA EEKSYTDGRRENPCEFNRGYRDYTP L P VG V T K E L E Y F P E T D K V W I E I G E T E G T F I V D S V ELLLMEE

FIG. 12